Received: 24 July 2018 | Revised: 17 October 2018 | Accepted: 26 October 2018

DOI: 10.1111/mec.14942

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data

Amanda Xuereb[1] | Christopher M. Kimber[2] | Janelle M. R. Curtis[3] | Louis Bernatchez[4] | Marie-Josée Fortin[1]

[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

[2]AVIAN Behavioural Genomics and Physiology Group, Department of Physics, Chemistry, and Biology (IFM), Linköping University, Linköping, Sweden

[3]Pacific Biological Station, Ecosystem Sciences Division, Fisheries and Oceans Canada, Nanaimo, British Columbia, Canada

[4]Institut de Biologie Intégrative et des Systèmes, Université Laval, Québec, Québec, Canada

Correspondence
Amanda Xuereb, Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada.
Email: amanda.xuereb@mail.utoronto.ca

Funding information
Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: 5134, D3-460408-2014 and STPGP 430706-2012; Program for Aquaculture Regulatory Research; Ressources Aquatiques Québec

Abstract

Understanding the spatial scale of local adaptation and the factors associated with adaptive diversity are important objectives for ecology and evolutionary biology, and have significant implications for effective conservation and management of wild populations and natural resources. In this study, we used an environmental association analysis to identify important bioclimatic variables correlated with putatively adaptive genetic variation in a benthic marine invertebrate—the giant California sea cucumber (*Parastichopus californicus*)—spanning coastal British Columbia and south-eastern Alaska. We used a redundancy analysis (RDA) with 3,699 single nucleotide polymorphisms (SNPs) obtained using RAD sequencing to detect candidate markers associated with 11 bioclimatic variables, including sea bottom and surface conditions, across two spatial scales (entire study area and within subregions). At the broadest scale, RDA revealed 59 candidate SNPs, 86% of which were associated with mean bottom temperature. Similar patterns were identified when population structure was accounted for. Additive polygenic scores, which provide a measure of the cumulative signal across all candidate SNPs, were strongly correlated with mean bottom temperature, consistent with spatially varying selection across a thermal gradient. At a finer scale, 23 candidate SNPs were detected, primarily associated with surface salinity (26%) and bottom current velocity (17%). Our findings suggest that environmental variables may play a role as drivers of spatially varying selection for *P. californicus*. These results provide context for future studies to evaluate the genetic basis of local adaptation in *P. californicus* and help inform the relevant scales and environmental variables for in situ field studies of putative adaptive variation in marine invertebrates.

KEYWORDS
climate data, environmental association analysis, marine invertebrate, RAD-sequencing, redundancy analysis, seascape genomics

1 | INTRODUCTION

Heterogeneity in environmental conditions imposes differential selection pressures across space, potentially leading to the adaptation of populations to local environments (Kawecki & Ebert, 2004). In the marine environment, nearshore regions are characterized by steep gradients in temperature, salinity and other abiotic factors, due to persistent upwelling systems, warming by the air and sun, exposure to freshwater run-off, and pollution from coastal developments (Breitburg, Hondorp, Davias, & Diaz, 2009; Foreman, Callendar, Masson, Morrison, & Fine, 2014). These factors have the potential to maintain adaptive polymorphisms via spatially varying selection (Bradbury et al., 2010; Gagnaire, Normandeau, Côté, Hansen, & Bernatchez, 2012; Laporte et al., 2016). Evidence of local adaptation associated with environmental conditions in marine populations is growing, challenging the dogma that a high degree of gene flow precludes adaptive divergence in the ocean (Bernatchez, 2016; Sanford & Kelly, 2011). Empirical studies have identified key environmental predictors of adaptive genetic variation across environmental gradients in marine taxa including corals (Lundgren, Vera, Peplow, Manel, & van Oppen, 2013), eels (Babin, Gagnaire, Pavey, & Bernatchez, 2017; Gagnaire et al., 2012), sticklebacks (Guo, DeFaveri, Sotelo, Nair, & Merilä, 2015), lobsters (Benestan et al., 2016), and Pacific (Hecht, Matala, Hess, & Narum, 2015) and Atlantic (Jeffery et al., 2017) salmon.

Marine and coastal environments are subject to rapid changes in physical and chemical properties including temperature, salinity, water circulation, pH and oxygen concentration (Hoegh-Guldberg et al., 2014), with important consequences for the development and survival of organisms inhabiting these regions (e.g., Gobler, DePasquale, Griffith, & Baumann, 2014; Kroeker et al., 2013; O'Connor et al., 2007). The ability for marine species to adapt to novel environmental conditions is increasingly important as unprecedented rates of environmental change continue to threaten marine biodiversity (Hoegh-Guldberg, Poloczanska, Skirving, & Dove, 2017; Munday, Donelson, & Domingos, 2017). Thus, understanding the spatial scale of local adaptation and the factors associated with adaptive genetic variation is important for the effective conservation and management of wild populations and natural resources (Bradbury et al., 2010; Bernatchez et al., 2017; Nielsen, Beger, Henriques, Selkoe, & Heyden, 2017; von der Heyden, 2017) and predicting evolutionary responses to climate change and environmental disturbances (Bay et al., 2017). Quantifying spatial patterns of adaptive genetic variation has implications for the spatial management of marine populations, and is inherently central to the establishment of marine reserves that aim to protect genetic diversity and promote resilience to environmental change (von der Heyden, 2017).

In the last decade, several analytical methods have been developed to detect putatively adaptive loci from genomic data sets, making it possible to assess patterns of adaptive genetic variation in wild populations (Jensen, Foll, & Bernatchez, 2016; Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015; Schoville et al., 2012). Differentiation-based detections of statistical outliers are now commonplace, whereby extreme levels of locus-specific population genetic differentiation (e.g., F_{ST}) indicate that differentiation may be driven by adaptive rather than neutral demographic or historical processes (Narum & Hess, 2011). Additionally, environmental association analyses (EAAs) directly associate allele frequencies and environmental conditions hypothesized to influence local adaptation to not only detect genetic variants putatively under selection, but also to characterize the environmental conditions contributing to adaptive genetic variation (Joost et al., 2007; Rellstab et al., 2015; Schoville et al., 2012; Sork et al., 2013). EAAs are especially promising because they are better able to detect relatively weak signals of selection compared to methods based on population differentiation (De Mita et al., 2013; de Villemereuil, Frichot, Bazin, François, & Gaggiotti, 2014; Forester, Lasky, Wagner, & Urban, 2018). In particular, multivariate EAA methods model the effect of a suite of environmental predictors on a large number of genetic loci simultaneously, thus minimizing the number of statistical tests. Furthermore, multivariate EAA methods are well suited to detect weak multilocus responses to environmental conditions by modelling the covariance of loci in response to environmental conditions (Bourret, Dionne, Kent, Lien, & Bernatchez, 2013; Forester et al., 2018; Rellstab et al., 2015).

In this study, we investigated the influence of bioclimatic factors on putatively adaptive genetic variation in a benthic marine invertebrate—the giant California sea cucumber (*Parastichopus californicus*)—in the northeastern Pacific Ocean. *P. californicus* undergoes a bipartite life cycle: a dispersive pelagic larval stage with a relatively long pelagic larval duration (up to 120 days; Lambert, 1997) and a relatively sedentary benthic adult stage that occupies rocky, sandy and algae-covered substrates in nearshore regions (Hamel & Mercier 2008). Commercial harvesting of this species occurs throughout its distribution along the Pacific coast of North America. In Canada, commercial harvesting of *P. californicus* is managed as a dive-only rotational fishery (DFO, 2014). Though sea cucumbers are not farmed in Canada, there is an interest to establish aquaculture facilities for *P. californicus* in British Columbia (DFO, 2016). As such, understanding the role of environmental factors as drivers of natural selection in this species could help inform management-based decisions that aim to protect adaptive potential for the future. Previous work on this species in the same region identified two distinct genetic groups using multiple approaches for evaluating population genetic structure (Xuereb et al., 2018). These groups corresponded to regions north and south of Queen Charlotte Sound at the northern tip of Vancouver Island (Figure 1), which corroborated patterns of population genetic structure observed in other marine species in the same region, including the rosethorn rockfish (*Sebastes helvomaculatus*; Rocha-Olivares & Vetter, 1999) and the bat star (*Pisaster miniata*; Sunday, Popovic, Palen, Foreman, & Hart, 2014). However, the potential for environmental conditions to influence spatial patterns of genetic variation has not yet been evaluated. Coastal British Columbia is predicted to experience decreases in surface salinity as a result of increases in freshwater run-off (Foreman et al., 2014; Morrison, Callendar, Foreman, Masson, & Fine, 2014), and increases in surface temperature and intensification of surface currents (Foreman et al.,

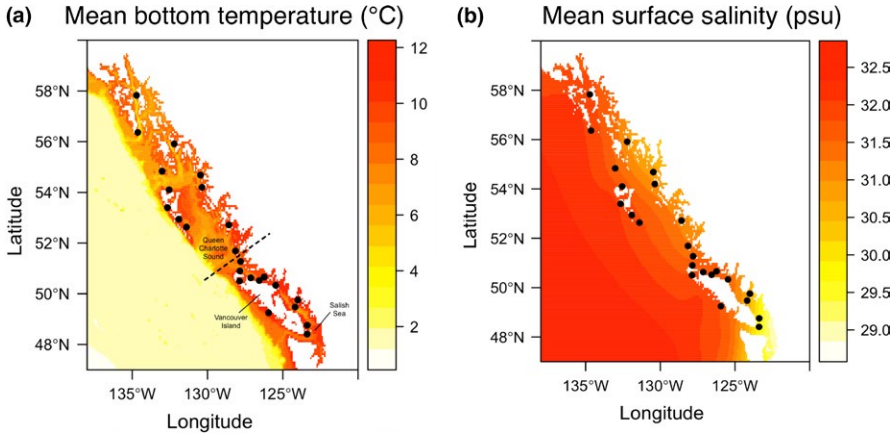

FIGURE 1 (a) Mean bottom temperature (data from BIO-ORACLE; Tyberghein et al., 2012) and (b) mean surface salinity (data from MARSPEC; Sbrocco & Barber, 2013) in the coastal region of British Columbia and southeastern Alaska; black dots indicate *P. californicus* sampling locations. The dashed line in (a) indicates the location of the genetic break identified in Xuereb et al. (2018) [Colour figure can be viewed at wileyonlinelibrary.com]

2014). Changes in oceanic heat content and thermal stratification are likely to influence subsurface conditions in the future as well (Abraham et al., 2013), potentially impacting both the pelagic and benthic life stages of *P. californicus* and other marine invertebrates.

Our main objective was to identify the environmental variables contributing to spatial patterns of putatively adaptive genetic variation across sampled locations. We tested for associations between allele frequencies at single nucleotide polymorphism (SNP) loci derived from restriction-site associated DNA sequencing (RADseq) and bioclimatic factors predicted to influence local adaptation of sea cucumbers using a multivariate constrained ordination EAA approach over two spatial scales: (a) across the entire sampled geographic area and (b) within regional groups. We considered candidate loci potentially under selection based on the loadings of SNPs in ordination space, and identified the environmental variables most strongly correlated with candidate loci. We then evaluated how candidate loci collectively varied with the best-associated environmental conditions across our sampling locations.

2 | METHODS

2.1 | Sampling and laboratory methods

The analyses in this study were performed on the same data presented in Xuereb et al. (2018). Whereas the previous study retained only those SNPs identified as putatively neutral for analyses of population structure and connectivity, we included the full set of filtered SNPs (3,699 SNPs) for analyses of adaptive genetic variation in this study. Sample collections, sequencing methods and filtering steps are described here in brief but we refer to Xuereb et al. (2018) for full details.

Tissue samples were collected from adult *Parastichopus californicus* by SCUBA from 24 sampling locations along the coastline of British Columbia (BC) and southeastern Alaska (Figure 1), with the number of individual specimens (spike clips) collected from each location ranging between 30 and 41. These sites capture latitudinal variability along a linear coastline as well as conditions within basins between Vancouver Island and the BC mainland and in more remote regions (Figure 1). Whole genomic DNA was extracted from individual tissue samples using the DNeasy spin column protocol (QIAGEN, Toronto, ON, Canada). Libraries were prepared using the RAD sequencing protocol of Poland, Brown, Sorrells, and Jannink (2012), and single-end sequencing was performed on the Ion Proton™ platform (Life Technologies, Grand Island, NY) at the core sequencing facility at the Institut de Biologie Intégrative et des Systèmes at Université Laval (Québec, Canada). Raw reads were aligned to the genome of a closely related species (*Parastichopus parvimensis*; Cameron, Samanta, Yuan, He, & Davidson, 2009) and a catalogue of putative loci was built with STACKS version 1.4.4 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013), allowing for a maximum of three mismatches between loci (n = 3). Initial filtering steps performed in STACKS included retaining SNPs with a minimum stack depth of four (m = 4), present in at least 16 sampling locations (one-third of the total number of locations), and found in at least 70% of individuals within each sampling location. We subsequently removed loci exhibiting a deficiency or excess of heterozygosity (H_o > 0.6 in at least one site; F_{IS}>0.5 or <−0.5), as well as loci with either a minor allele frequency (MAF) less than 0.01 across all sites or less than 0.1 within at least one site. We also excluded one locus per pair of loci in high linkage disequilibrium (LD) (R^2 > 0.8) from the full set of markers, as well as samples with more than 30% missing data. See Xuereb et al. (2018) for a full description of the parameters used to filter the data set and the number of loci removed after each filtering step.

2.2 | Environmental predictors

We considered a total of 14 environmental variables including temperature, salinity, current velocity, dissolved oxygen concentration and chlorophyll concentration at each of the 24 sampling locations (Supporting Information Table S1) as potential drivers of spatially varying selection. We used monthly surface temperature data (°C) between 2002 and 2010 and monthly surface salinity data (psu) between 1955 and 2006 from the MARSPEC (Sbrocco & Barber, 2013) database at a resolution of 1 km (originally derived from the World Oceans Atlas 2009 (Antonov et al., 2010) and Aqua-MODIS (http://oceancolor.gsfc.nasa.gov/), respectively). We included the mean, minimum and maximum monthly measurements for both surface temperature and salinity, as selection may be influenced by both

average and extreme climatic conditions. Using a principal components analysis (PCA), sea surface salinity and sea surface temperature (mean, minimum and maximum) were reduced to three principal components (PCs) that explained 97.5% of the total variation: PC1 was negatively correlated with sea surface salinity; PC2 was positively correlated with mean and maximum sea surface temperature; PC3 was negatively correlated with minimum surface temperature (Table 1). These PC axes were subsequently used as predictor variables representing sea surface salinity (PC1) and sea surface temperature (PC2 and PC3) in the constrained ordinations (see below). As environmental conditions at the bottom can also influence local adaptation in populations of benthic organisms, we additionally included data for bioclimatic variables representing conditions at bottom depth, which were obtained from the BIO-ORACLE database (Tyberghein et al., 2012; Supporting Information Table S2). These data were assembled at a coarser resolution (9.2 km) compared to the MARSPEC data. We evaluated the pairwise correlation values between all BIO-ORACLE predictors and excluded variables demonstrating high correlation with other bioclimatic variables. Significance of the correlations was assessed based on *p*-values corrected for multiple tests using the Benjamini-Hochberg (BH) method ($p < 0.05$) (Benjamini & Hochberg, 1994). Lastly, we assessed multicollinearity between all predictor variables prior to performing constrained ordinations (see below), using the variance inflation factor (VIF). Predictor variables with VIF > 10 were excluded, and all predictors were scaled and centred prior to analyses. MARSPEC and BIO-ORACLE data were obtained using the SDMPREDICTORS version 0.2.5 package (Bosch, 2017) in R (R Core Team 2016).

2.3 | Environmental association analysis: constrained ordinations

We performed a redundancy analysis (RDA) to detect candidate adaptive loci exhibiting strong associations with the environmental variables hypothesized to influence selection. RDA is an extension of linear regression in which both the predictor and the response variables are multivariate. This approach performs a PCA on the response table (here, the matrix of allele frequencies) while constraining the PCA axes as linear combinations of the predictor variables (i.e., environmental variables). Allele frequencies were Hellinger transformed prior to running the RDA (Legendre & Gallagher, 2001). All RDAs were performed using the *rda* function in the VEGAN 2.4-5 (Oksanen et al., 2017) package in R.

We estimated the proportion of variance in allele frequencies at 3,699 SNPs across all 24 sampling locations that could be explained by environmental predictors based on the adjusted R^2. We used an analysis of variance (ANOVA) with 1,000 permutations to evaluate the significance of the global RDA. Then, we identified candidate loci based on locus scores (i.e., the loading of each locus in ordination space) that were ±3 *SD* from the mean loading (following Forester et al., 2018) on the first two constrained ordination axes. We identified the environmental variables exhibiting the strongest associations with each candidate adaptive locus using a Pearson's correlation (*r*).

To further investigate the spatial distribution of putatively adaptive polymorphisms across our sampling locations, we evaluated the MAF at the candidate loci most strongly correlated with environmental variables within each of the 24 sampling locations. For this, we used an arbitrary cut-off of $r = 0.65$ to select candidate loci showing the strongest correlations with environment. For each locus that met this criterion, we performed a linear regression to assess the relationship between MAF and the value of the best-associated environmental variable across sampling locations. We used a BH *p*-value adjustment to correct for multiple tests (Benjamini & Hochberg, 1994).

In structured populations, genetic signatures of selection may be confounded with signatures of neutral (i.e., historical or demographic) processes that may be falsely interpreted as selection (Excoffier, Hofer, & Foll, 2009). Corrections for population structure are thus recommended to control for signals generated by neutral processes, although these corrections can also be overly conservative in some cases by inadvertently removing true signals of selection (Forester et al., 2018). Nonetheless, we performed a second RDA in which we account for population structure to determine its effect on detections of candidate loci. We used a spatial eigenfunction analysis to account for population structure following a similar approach to that described in Forester et al. (2018). We computed distance-based Moran's eigenvector maps (dbMEMs) based on the Euclidean distances between sampling locations, which decompose the spatial relationships among sampled sites into a set of spatial variables (Dray, Legendre, & Peres-Neto, 2006). We used the R package ADESPATIAL version 0.1-0 (Dray et al., 2017) to compute dbMEMs. First, a PCA was performed on Hellinger-transformed allele frequencies across all 3,699 loci and retained the first three PC axes based on visualization of a screeplot. We performed a RDA with the retained PC axes as the response variables and all dbMEMs as the explanatory variables, and subsequently used a forward selection procedure (Blanchet, Legendre, & Borcard, 2008) to identify significant dbMEM variables. The significant dbMEMs were then used as conditioning variables in a partial RDA and candidate SNPs after

TABLE 1 Factor loadings of the MARSPEC bioclimatic variables on the first three principal component axes with per cent of variance explained by each axis in parentheses

Bioclimatic variable	PCA1 (60%)	PCA2 (24%)	PCA3 (14%)
SSS annual mean	−0.493	0.267	−0.157
SSS monthly min	−0.486	0.282	−0.044
SSS monthly max	−0.481	0.279	−0.224
SST annual mean	0.368	0.566	−0.113
SST coldest ice free month	0.322	0.144	−0.836
SST warmest ice free month	0.226	0.656	0.460

Note. Loadings represent correlation coefficients between the variables and the PC axes.
SSS: sea surface salinity; SST: sea surface temperature.

correcting for population structure were detected using the same methods as described previously.

2.4 | Comparison with differentiation-based outlier detection

In addition to RDA, we used a population differentiation-based approach to detect candidate loci under selection to compare detections across methods. Specifically, we used Bayescan, which estimates the posterior probability of SNPs under selection based on F_{ST} values (Foll & Gaggiotti, 2008). This approach was used previously on the same data set to retain only putatively neutral markers for analyses of population genetic structure, using prior odds of 10,000, a q-value threshold of 0.01, and with 10,000 iterations and 200,000 burn-in steps (see Xuereb et al., 2018). Here, we retained the 55 SNPs (out of a total of 3,699 SNPs) identified as being under divergent selection to determine the proportion of candidate markers detected by both methods.

2.5 | Additive polygenic scores

We used an approach based on additive polygenic scores for each individual to evaluate the cumulative effect of all candidate loci in response to environmental conditions (Gagnaire & Gaggioti 2016), following Babin et al. (2017). We first assessed the extent of LD based on R^2 values between all pairs of loci identified as being under selection to ensure that candidate loci were not strongly linked. Polygenic scores were calculated by first identifying alleles across all candidate loci that were associated with increasing values of a given environmental variable (e.g., mean bottom temperature) based on the direction of correlation between allele frequencies and the environmental condition. A score was calculated for each individual sea cucumber by summing the total number of favoured alleles within a particular environment across all candidate loci. Then, we evaluated the relationship between individual additive polygenic scores and environmental variables independently using both a linear and a quadratic model, and determined the best-fit model based on the lowest Akaike information criterion (AIC) value. We performed the polygenic score analysis separately using the candidate loci detected without correction for population structure (by both RDA and Bayescan) and with correction for population structure (by partial RDA), and compared the two approaches.

2.6 | Candidate SNPs under selection at a finer spatial scale

We performed a second RDA using the same methods as described above within north and south regional areas independently (with no correction for population structure) to determine whether selection pressures vary between whole-coast and within-region scales. These two areas consisted each of 12 sampling sites located north and south of Queen Charlotte Sound identified as belonging to two distinct genetic clusters in Xuereb et al. (2018) (Figure 1).

3 | RESULTS

3.1 | Sequencing and marker filtering

On average, 2.75 million raw reads per sample were obtained from RAD sequencing and aligned with the *Parastichopus parvimensis* genome. A total of 94,842 SNPs were retained from the catalogue of ~1.81 million putative loci. Subsequent filtering steps based on minimum presence, observed heterozygosity, and local and global minor allele frequencies retained 4,340 SNPs. After excluding one locus from each pair in LD ($R^2 > 0.8$), a total of 3,699 SNPs were retained, in 717 individuals across the 24 sampling locations, for subsequent analyses. Further details are provided in Xuereb et al. (2018).

3.2 | Environmental predictors

Several significant correlations were observed between pairs of the BIO-ORACLE environmental variables (Supporting Information Table S2). As such, we opted to exclude three variables: "bottom dissolved oxygen concentration," "bottom chlorophyll concentration" and "maximum temperature at the bottom" from our set of predictors. All of the remaining predictor variables, including PC axes representing surface salinity and surface temperature, had a VIF < 10, both at the broad scale (i.e., among all 24 sites) and within the south region, indicating no issues of multicollinearity among the predictors (Table 2). We excluded bottom current velocity from the analysis within the north region as it had a VIF > 10.

3.3 | Environmental association analysis: constrained ordinations

The RDA with all 3,699 SNPs including all 24 sampling locations was globally significant (ANOVA $F_{(8,15)} = 1.26$, $p = 0.001$) and explained about 8% of the variance (adjusted $R^2 = 0.083$) (Figure 2a). Only the first RDA axis was significant ($p = 0.002$). However, we considered

TABLE 2 Variance inflation factor (VIF) for predictor variables included in the constrained ordinations at the broad scale (VIF_{all}) and among the sampling sites within the south (VIF_{south}) and north (VIF_{north}) groups, excluding variables with VIF>10

Environmental variables	VIF_{all}	VIF_{south}	VIF_{north}
PC1	2.06	2.19	1.63
PC2	1.19	1.98	2.47
PC3	1.17	1.81	1.45
Current velocity (B)	2.13	1.97	—
Mean temperature (B)	1.40	3.65	1.26
Minimum temperature (B)	2.66	2.43	2.03
Mean salinity (B)	1.30	2.85	2.07
Current velocity (S)	1.53	1.96	2.19

Note. (B) = bottom and (S) = surface.

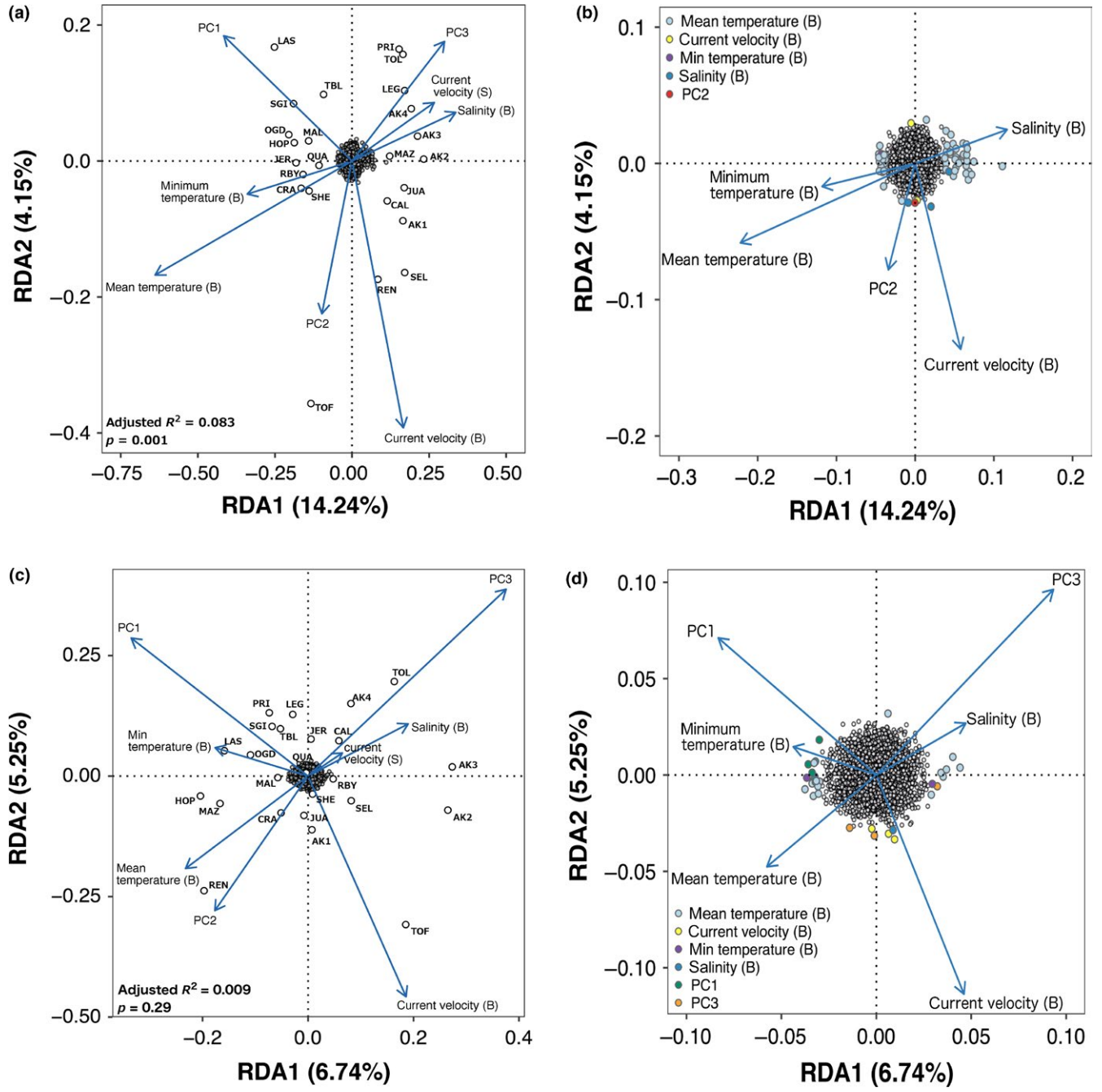

FIGURE 2 Redundancy analysis (RDA) performed with 3,699 SNPs (grey filled circles in the centre) and eight environmental variables (blue arrows) on the first two constrained ordination axes (a) with no correction for population structure and (c) with correction for population structure by conditioning on significant dbMEMs. Candidate loci detected based on locus scores ±3 *SD* from the mean loading on each RDA axis are shown in zoomed-in plots for (b) the uncorrected RDA and (d) the RDA with correction for population structure, and coloured by the most highly correlated environmental explanatory variable. SNPs not identified as candidates are shown in grey; blue arrows represent the environmental predictors correlated with candidates loci [Colour figure can be viewed at wileyonlinelibrary.com]

candidate loci on the first two constrained canonical axes, which explained 14.2% and 4.15% of the genetic variation, respectively. Based on locus scores that were ±3 *SD* from the mean loadings, 59 loci were identified as candidates (51 candidate loci were detected on RDA axis 1 and 8 candidate loci were detected on RDA axis 2; Figure 2b). The majority of candidate loci (51; 86%) were associated with mean bottom temperature. Of the remaining candidate loci, 3 (5%) were associated with mean bottom salinity, 3 (5%) were associated with bottom current velocity, 1 (2%) was associated with minimum bottom temperature, and 1 (2%) was associated with mean and maximum surface temperature (PC2).

Overall, sixteen out of the 59 candidate loci identified by RDA exhibited correlation coefficients (*r*) greater than 0.65 (either positively or negatively correlated) with environment, all of which

were associated with mean bottom temperature (Supporting Information Table S3). In all but one of these 16 loci, minor allele frequencies were negatively correlated with mean bottom temperature, indicating that MAF at these candidate loci is lower in sampling locations with higher bottom temperatures (Figure 3). The relationship between MAF and temperature at sampling locations was significant for all loci after correcting for multiple tests (Table 3).

Two significant spatial variables (dbMEM1 and dbMEM2) were identified using forward selection and were retained as conditioning variables in a partial RDA to detect candidate loci while correcting for broad-scale population structure. The partial RDA was not significant overall (ANOVA $F_{(8,13)} = 1.03$, adjusted $R^2 = 0.009$, $p = 0.29$), likely as a result of the large number of neutral SNPs contributing to the total genetic variation (Figure 2c). Nevertheless, we identified candidate SNPs in this partial model using the same threshold as above (±3 *SD* from the mean loading). Using this approach, 33 outliers were detected on the first two constrained axes: 26 on the first RDA axis and seven on the second RDA axis (Figure 2d). Of these outlier loci, 17 (52%) were associated with mean bottom temperature; 4 (12%) each with surface salinity (PC1) and minimum surface temperature (PC3); 3 (9%) each with bottom current velocity and bottom salinity; and 2 (6%) with minimum bottom temperature. Considering only the candidate markers that were associated with mean bottom temperature, as these made up the majority of candidates, 14 out of the 17 outlier SNPs detected in the partial RDA were also detected without any correction for population genetic structure.

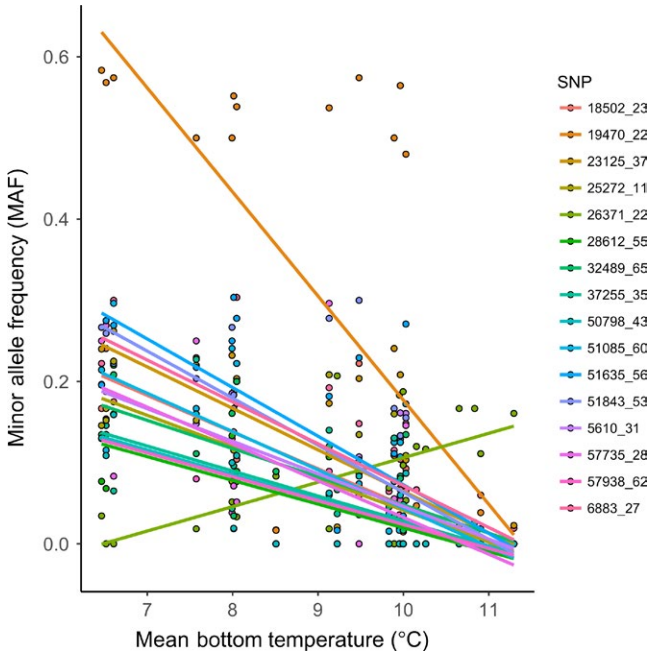

FIGURE 3 Minor allele frequency (MAF) as a function of mean bottom temperature (°C) for each of the 16 SNPs exhibiting the strongest correlation ($r > 0.65$) with bottom temperature. Solid lines represent the fit from a linear regression for each candidate SNP [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | Comparison with differentiation-based outlier detection

Of the 59 candidate SNPs under selection identified by RDA and the 55 candidate SNPs identified by Bayescan, a total of 43 detections overlapped between the two methods. Thus, across both methods, a combined total of 71 unique candidate SNPs were detected at the broadest spatial scale.

3.5 | Additive polygenic scores

Pairwise LD between all detected candidate loci was weak on average (mean $R^2 = 0.04$) with 96% of pairwise R^2 values below 0.2 (all pairwise R^2 values are shown in Supporting Information Table S4), suggesting that candidate loci are not in strong LD. We first evaluated the relationship between additive polygenic scores for each individual sea cucumber across the 71 putatively adaptive loci detected by both methods described above, and mean bottom temperature. Additive polygenic scores increased significantly with increasing mean bottom temperature (linear model: $R^2 = 0.251$, $p < 2.2 \times 10^{-16}$) (Figure 4a). The quadratic model had only a slightly lower AIC score (quadratic model: AIC = 5,393.34; linear model: AIC = 5,394.20) and explained a similar proportion of variation in allele frequencies at candidate markers (quadratic model: $R^2 = 0.253$, $p < 2.2 \times 10^{-16}$) compared to a linear model (Supporting Information Figure S1a). We performed the analysis again using polygenic scores calculated over the 33 candidate loci that were detected by the partial RDA after correcting for population structure and observed a similar significant positive association

TABLE 3 Linear regression statistics for each of the candidate loci showing strong correlation ($r > 0.65$) with mean bottom temperature

SNP ID	Adjusted R^2	*F*	*P*	P_{adj}
18502_23	0.49	23.15	<0.0001	0.0002
19470_22	0.44	18.85	0.0003	0.0003
23125_37	0.45	20.18	0.0002	0.0002
25272_11	0.46	20.54	0.0002	0.0002
26371_22	0.37	14.54	0.0010	0.0010
28612_55	0.52	25.49	<0.0001	0.0001
32489_65	0.59	34.07	<0.0001	<0.0001
37255_35	0.43	18.42	0.0003	0.0003
50798_43	0.56	30.85	<0.0001	<0.0001
51085_60	0.49	23.51	<0.0001	0.0002
51635_56	0.52	25.51	<0.0001	0.0001
51843_53	0.47	21.69	0.0001	0.0002
5610_31	0.47	21.14	0.0001	0.0002
57735_28	0.40	16.28	0.0006	0.0006
57938_62	0.56	30.02	<0.0001	<0.0001
6883_27	0.61	36.64	<0.0001	<0.0001

Note. Adjusted *P*-values (P_{adj}) are based on a Benjamini-Hochberg correction for multiple tests.

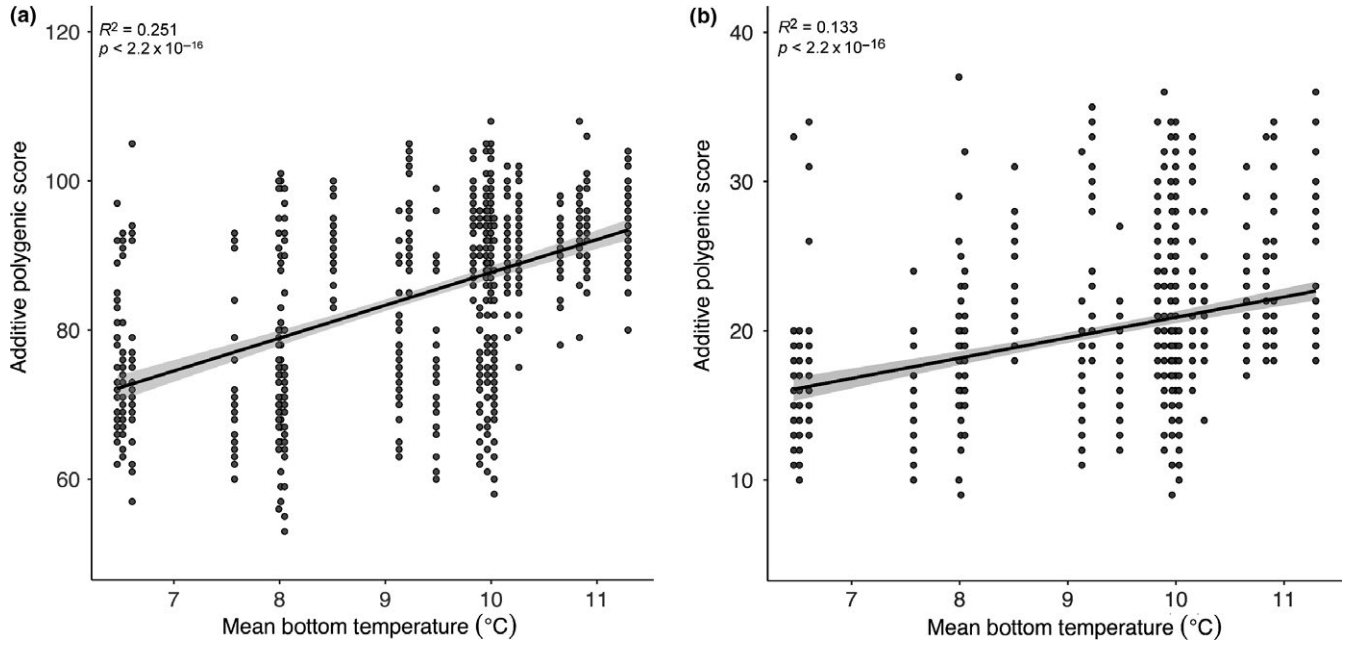

FIGURE 4 Relationship between individual additive polygenic scores calculated (a) across 71 candidate SNPs identified by RDA with no correction for population structure and Bayescan and (b) across 33 candidate SNPs identified by RDA after correcting for population structure, and the mean bottom temperature across sampling locations. The solid line represents the regression line from the linear model, and the shaded area represents the 95% confidence interval. Each dot represents an individual sea cucumber [Colour figure can be viewed at wileyonlinelibrary.com]

between polygenic scores and mean bottom temperature, albeit with lower statistical support given the reduced number of loci (R^2 = 0.133, $p < 2.2 \times 10^{-16}$) (Figure 4b and Supporting Information Figure S1b). AIC scores for the quadratic and linear models were similar (quadratic model: AIC = 5,393.34; linear model: 5,394.20).

3.6 | Candidate SNPs under selection at a finer spatial scale

Within the north region, the RDA was not significant (ANOVA $F_{(7,4)}$ = 0.95, p = 0.93) and the proportion of variance explained by the predictor variables was negligible (adjusted R^2 = −0.034); thus, we did not consider candidate loci and associations with environmental factors within this region. In the south region, the proportion of total genetic variance explained by the predictor variables was similar to that observed for the total data set (adjusted R^2 = 0.07), and the global RDA was significant (ANOVA $F_{(8,3)}$ = 1.11, p = 0.017) (Figure 5a). As with the broad coastal scale, we considered candidate loci on the first two constrained axes, which explained 13.4% and 10.2% of the genetic variance, respectively. We detected 23 candidate loci: 11 on the first RDA axis and 12 on the second RDA axis (Figure 5b). Of these candidate loci, 6 (26%) were associated with surface salinity (PC1), 4 (17%) were associated with bottom current velocity, 3 (13%) were associated with minimum bottom temperature, 3 (13%) were associated with mean bottom temperature, 3 (13%) were associated with bottom salinity, 2 (7%) were associated with PC3 (minimum surface temperature), 1 (4%) was associated with PC2 (mean and maximum bottom temperature), and 1 (4%) was associated with surface current velocity. At the regional scale, locus-specific F_{ST} values were very low (maximum F_{ST} = 0.005) and Bayescan did not detect F_{ST} outliers under divergent selection at this scale with a q-value threshold of 0.01.

4 | DISCUSSION

In this study, we investigated the influence of environmental features as potential drivers of adaptive divergence in a benthic marine species, the giant California sea cucumber (*Parastichopus californicus*). The primary question we asked was: What are the environmental conditions driving differentiation at putatively adaptive genetic loci? Using a multivariate EAA, we identified a subset of candidate loci from a data set of 3,699 SNPs derived from RAD sequencing exhibiting associations with bioclimatic variables hypothesized to influence spatially varying selection in benthic marine organisms. We also used an approach based on additive polygenic scores to assess the relationship between candidate markers collectively and mean bottom temperature, which was identified as an important predictor of genetic variation at putative loci under selection, across sampling locations. Moreover, different genotype-environment associations were identified within the south region, compared to the entire geographic range included in our analyses, implying that selection pressures may differ across spatial scales. The results of these analyses imply that environmental conditions, especially bottom temperature and/or surface salinity, may play important roles as drivers of spatial patterns of putative adaptive genetic variation that could influence local adaptation.

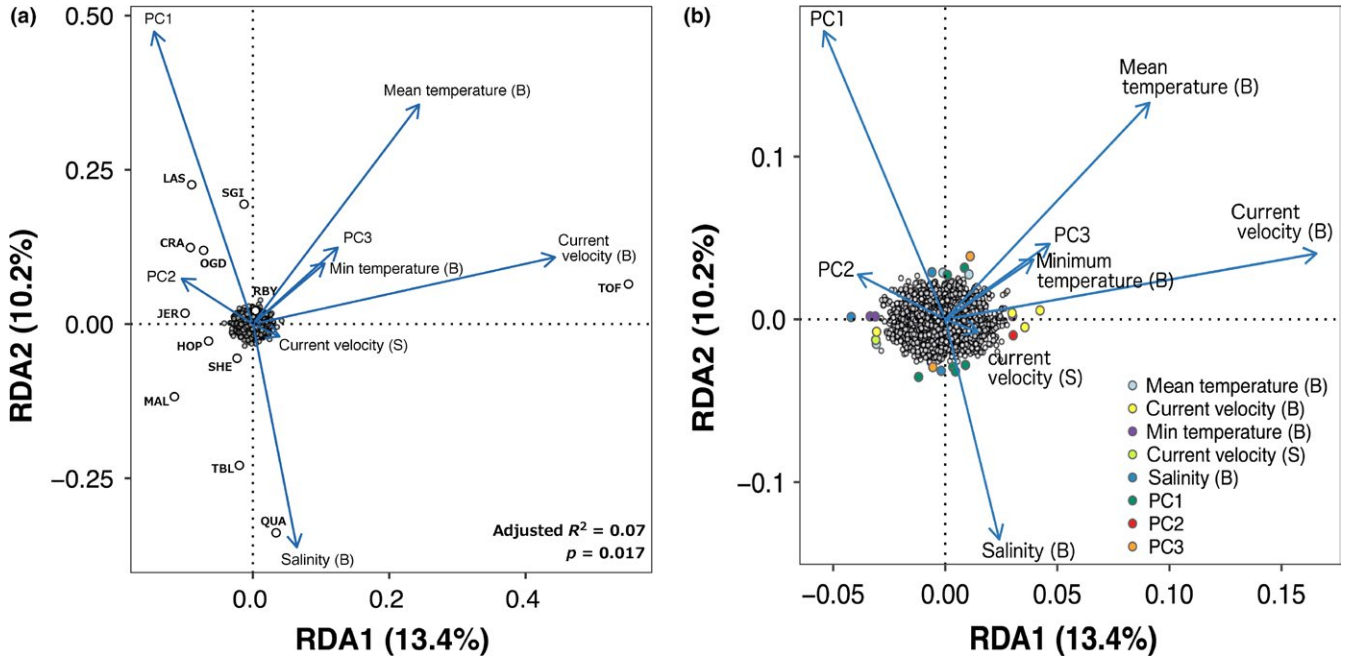

FIGURE 5 (a) Redundancy analysis (RDA) performed within the south region with 3,699 SNPs (grey filled circles in the centre) and eight environmental variables (blue arrows) on the first two constrained ordination axes; (b) zoomed-in plot showing candidate loci detected based on locus scores ±3 *SD* from the mean loading on each RDA axis coloured by the most highly correlated environmental explanatory variable. In (b), SNPs not identified as candidates are shown in grey; blue arrows represent the environmental predictors correlated with candidates loci [Colour figure can be viewed at wileyonlinelibrary.com]

Similar to other studies using EAA approaches to detect candidate loci under selection in marine systems (e.g., Bay & Palumbi, 2014; Benestan et al., 2016; De Wit & Palumbi, 2013; van Wyngaarden et al.., 2017), we identified a subset of SNPs showing strong associations with environmental factors. A larger set of candidate loci (59) was detected across all 24 sampling locations compared to within the south region alone (23) using RDA. In our study, restricted gene flow associated with ocean current circulation between the north and south subregions (Xuereb et al., 2018) may facilitate divergent selection at this broad spatial scale. Though this previous study found significant, albeit weak substructure within subregions (AMOVA, F_{ST} within regions = 0.002, p = 0.001), the detection of outlier SNPs within the south region independently suggests that adaptive differentiation may indeed occur in the presence of considerable gene flow (Yeaman & Otto, 2011).

4.1 | Environmental drivers of adaptive differentiation

Our results suggest that sea bottom temperature is an important predictor of genetic variation at candidate loci over a broad spatial scale and thus may be a potential driver of spatially varying selection for *P. californicus*. Sea bottom temperature represents the thermal environment experienced by the benthic life stage and may thus be a selective agent acting on settlers and/or adults. Indeed, other studies have demonstrated that temperature is a significant determinant of adaptive differentiation among populations or of clinal patterns of adaptive genetic structure in other marine species (Stanley et al., 2018), such as Atlantic cod exhibiting parallel clines in variation on either side of the Atlantic Ocean (Bradbury et al., 2010), as well as in other marine invertebrates including purple sea urchins (*Strongylocentrotus purpuratus*) spanning a latitudinal gradient from the Pacific coast of Canada to Baja California (Pespeni & Palumbi, 2013), reef-building corals (*Acropora hyacinthus*) occupying pools with different thermal conditions in the south Pacific Ocean (Bay & Palumbi, 2014; Palumbi, Barshis, Traylor-Knowles, & Bay, 2014) and American lobster inhabiting spatially varying temperature regimes in the Northwest Atlantic ocean (Benestan et al., 2016). It is possible that temperature may not be the direct causative agent of selection, but instead adaptation may be directly attributable to other variables that are correlated with temperature. Nevertheless, variability in water temperature is known to drive divergence in marine systems by selecting on thermal tolerance at multiple life history stages, affecting traits related to growth (Yanick, Heath, & Heath, 2003), survival (Kuo & Sanford, 2009; Osovitz & Hoffman, 2005; Palumbi et al., 2014) and reproduction (Pardo & Johnson, 2005).

The strong positive correlation observed between individual polygenic scores calculated across all candidate loci and mean bottom temperature is consistent with spatially varying selection across a temperature gradient, where different alleles are maintained in different thermal environments. A quadratic model did not describe the relationship between polygenic scores and temperature considerably better than a linear model, implying that the strength

of selection is relatively constant across the temperature gradient, though the number of candidate loci may limit our capacity to detect minor differences in selection strength. It is important to note that the interpretation of the role of polygenic selection may be somewhat biased since we cannot completely dismiss the possibility that some candidate loci are physically linked without a reference genome. Yet, the lack of evidence for strong statistical LD is suggestive that these markers are likely not physically linked. We also observed that the major allele showed evidence of approaching fixation at high temperatures for 15 of the 16 loci most strongly associated with bottom temperature (MAF close to 0), with the minor allele segregating at low to intermediate frequencies in cold temperatures; only 1 locus showed the opposite pattern. This could imply that, at least for the loci most strongly associated with bottom temperature, the strength of selection may be higher in warmer temperatures. In contrast, higher levels of polymorphism in cooler sites may be maintained via several mechanisms, including spatial and/or temporal balancing selection (Bergland, Behrman, O'Brien, Schmidt, & Petrov, 2014; Bernatchez, 2016). Better characterization of the relative strength of selection across the temperature gradient would benefit from a greater number of candidate loci, and disentangling the selective agents driving adaptive differentiation requires experimental validation of allelic effects, combined with functional information about the phenotypic impact of loci involved (e.g., annotated genomic resources). Additionally, increased knowledge of the demographic and environmental patterns over time (e.g., joint time-series data on temperature and allele frequency dynamics) would improve our understanding of changes in allele frequencies in response to changing environmental conditions.

Within the south region, surface salinity and bottom current velocity were the most strongly associated environmental predictors of genetic variation at candidate loci, given both the number of candidate loci associated with environmental predictors and the strength of correlations. Surface salinity may be influenced by freshwater input from coastal regions, which could produce a selective gradient over relatively small spatial scales (Bible & Sanford, 2016). In this particular region, the Fraser River discharges freshwater into the Salish Sea, potentially leading to localized decreases in sea surface salinity, whereas salinity is more homogeneous in the north area (Figure 1b). Echinoderms in particular may be sensitive to salinity stress as a result of their water vascular system and poor ion regulation capabilities (Binyon, 1972; Russell, 2013; Stickle & Diehl, 1987). Additionally, as reduced ocean circulation can generate hypoxic conditions in the benthos (Matear et al., 2000G), variability in current velocities may result in spatial variation in dissolved oxygen concentrations. These factors could potentially lead to adaptive divergence in physiological traits to cope with stressful conditions and should be investigated further.

4.2 | Correcting for population structure

An important concern in statistical approaches that identify candidate loci under selection is: To what extent do detected loci reflect a true adaptive signal vs. one caused primarily by demographic or historical processes? On the one hand, methods that correct for population genetic structure can help eliminate potentially spurious detections of candidate loci with allele frequency patterns that resemble selection but are caused by neutral processes (de Villemereuil et al., 2014; Excoffier et al., 2009). On the other hand, corrections for spatial structure may be too conservative, potentially removing true adaptive signals and resulting in an overall loss of power to detect loci under selection (Forester et al., 2018). This is because selection also plays a fundamental role in generating spatial patterns of population genetic structure (Charlesworth, Nordborg, & Charlesworth, 1997), and when population structure is confounded with environmental variables driving local adaptation, correcting for population structure can effectively eliminate the signal of selection one is aiming to identify (Forester et al., 2018; Yeaman et al., 2016). Indeed, in a simulation-based study that tested the performance of EAA methods, Forester et al. (2018) found that RDAs that did not correct for population structure actually performed better than those that included spatial variables. Specifically, they showed that false-positive rates (FPRs) were elevated and true-positive rates (TPRs) decreased considerably when correcting for population genetic structure, whereas RDAs that did not include such a correction exhibited low FPRs (and high TPRs) even when simulated demographic scenarios represented refugial expansions that covaried with the environmental gradient.

In our study region, previous work identified significant population structure, splitting north and south regional groups along the BC continental shelf (Xuereb et al., 2018), and thus, disentangling true environmental effects from historical or demographic processes is challenging. However, results obtained from an analysis with no correction for population structure are comparable to those obtained when using a more conservative approach that corrects for population structure (although with fewer candidate loci overall and a slightly weaker relationship between candidate loci and mean bottom temperature when the correction is applied). This implies that correcting for population structure may indeed exclude loci that are potentially under selection. Nonetheless, the observation that bottom temperature remains a notable predictor of genetic variation across candidate loci even after accounting for population structure lends support for an effect of environmental selection driven by bottom temperature. As such, bottom temperature has likely contributed to driving divergence between *P. californicus* populations and may continue to do so as ocean temperatures shift.

4.3 | Limitations and future directions

The lack of genomic resources, such as annotated reference genomes, for *P. californicus* and echinoderms more broadly, presents a challenge for identifying candidate genes underlying observed relationships. The availability of an annotated reference genome would allow matching of potential candidate loci to the genomic regions under selection and gives insights into the specific genes involved in local adaptation (Manel et al., 2016). The vast majority

of nonmodel systems will not have such resources readily available and the cost and time to assemble a sufficiently high-quality reference genome will likely be prohibitive in most cases. Annotated reference genomes from related species can also be used to identify candidate genes. However, the *Parastichopus parvimensis* genome that we used to align raw sequencing reads is not yet annotated (Cameron et al., 2009). Instead, we attempted to identify loci under selection using the purple sea urchin genome (*Strongylocentrotus purpuratus*; Sea Urchin Genome Sequencing Consortium, 2006), but this resulted in low-scoring alignments with RAD sequences containing candidate SNPs. As genomic resources become more widely available for a greater number of organisms, future work should aim to identify the genomic basis of loci putatively under selection. Moreover, the observed associations between allele frequencies and bioclimatic factors presented in this study can serve as hypotheses for further investigation into the causal relationship between environmental conditions and adaptive variation in *P. californicus* and other benthic marine invertebrate species with experimental evaluations of fitness differences (Savolainen, Lascoux, & Merilä, 2013).

A second limitation is that reduced representation sequencing approaches, like RADseq, sample a relatively small proportion of the entire genome and many loci under selection may be missed due to the sparse coverage of the genome (Lowry et al., 2017). Given the low density of markers sampled here, we did not aim to elucidate the genomic mechanisms underlying local adaptation. Rather, we focused on identifying the environmental features that may drive spatial patterns of selection and should be investigated further. The results of this study could potentially provide guidance for future projects to capture the geographic distribution of adaptive genetic variation while minimizing costs associated with more intensive genomic sampling per individual (e.g., whole-genome sequencing; see Fuentes-Pardo & Ruzzante, 2017).

4.4 | Implications for management and conservation

The ability to detect putatively adaptive genetic variation in wild populations has increased attention towards integrating measures of adaptive genetic variation into conservation decision-making processes (Flanagan, Forester, Latch, Aitken, & Hoban, 2017; Funk, McKay, Hohenlohe, & Allendorf, 2012), and recent studies have discussed the importance of incorporating estimates of adaptive potential into marine reserve design (von der Heyden, 2017; Nielsen et al., 2017). *P. californicus* is a commercially exploited species, and an understanding of the spatial patterns of adaptive genetic differentiation can inform management efforts to ensure the sustainability of the sea cucumber fishery in the future. While sea cucumbers are not presently farmed in Canada, there is an interest in the development of aquaculture programmes for *P. californicus* in British Columbia (DFO, 2014). Improved knowledge about local adaptation has important implications for effectively managing breeding programmes and informing the foundation of aquaculture broodstock (Do Prado et al., 2018). Further insight into the extent to which populations are locally adapted to environmental conditions will be important for ensuring that potentially maladapted genotypes are not introduced into wild populations (Conover, 1998). Moreover, understanding the spatial distribution of putatively adaptive genetic variation can inform the selection of specific sites for protection within marine reserves to maintain adaptive potential and evolutionary resilience of wild populations in the face of environmental change (von der Heyden, 2017). In some cases, populations that are already locally adapted to stressful or extreme environmental conditions might be important sources of "pre-adapted" alleles that can enhance the resistance of other populations to future environmental change (e.g., Bay & Palumbi, 2014; Golbuu, Gouezo, Kurihara, Rehm, & Wolanski, 2016). Our study did not identify unique locally adapted populations, but rather demonstrated the potential for environmental selection to shape the distribution of adaptive genetic variation across space. This can also have important implications for prioritizing sites for protection, where the level of adaptive genetic variation could be an indicator of the evolutionary resilience of populations (Bonin, Nicole, Pompanon, Miaud, & Taberlet, 2007; Sgrò, Lowe, & Hoffmann, 2011). In the coastal region of the northeastern Pacific, temperature and heat content are predicted to increase (Abraham et al., 2013; Foreman et al., 2014) and salinity is predicted to decrease (Foreman et al., 2014; Morrison et al., 2014) in the future. As such, the spatial patterns of genetic variability observed in this study can inform conservation planning decisions by identifying for protection populations containing higher levels of segregating polymorphisms associated with environmental conditions (e.g., temperature and salinity) that are prone to change in the future.

ACKNOWLEDGEMENTS

We gratefully acknowledge those involved in sample collections, especially Nicholas Duprey, Dan Curtis, Aaron Eger and Lynn Lee (Fisheries and Oceans Canada), Isabelle Côté (Simon Fraser University), Mike Donnellan, Kyle Hebert and Jeff Meucci (Alaska Department of Fish and Game). We thank Cécilia Hernandez, Brian Boyle and Gaetan Légaré (Université Laval) for laboratory assistance and sequencing, and Éric Normandeau for bioinformatics support. Finally, we also thank Brenna Forester and Nusha Keyghobadi for helpful feedback while drafting the manuscript, as well as Michael Hansen and three anonymous referees for comments on a previous version of the manuscript. We acknowledge support from the Program for Aquaculture Regulatory Research (PARR, DFO), Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic grant to M-J Fortin, J Curtis, L Bernatchez, I Côté, and F Guichard (#STPGP 430706-2012), NSERC Discovery Grant to M-J Fortin (#5134), the Canadian Research Chair in Genomics and Conservation of Aquatic Resources (L Bernatchez), Ressources Aquatiques Québec (RAQ) and NSERC Canada Graduate Scholarship to A Xuereb (#D3-460408-2014).

AUTHOR CONTRIBUTIONS

A.X., M.-J.F., J.M.R.C. and L.B. conceived and designed the study. A.X. analysed the data and wrote the manuscript, with support from C.M.K. All authors contributed to editing and revising the manuscript.

DATA ACCESSIBILITY

Raw de-multiplexed sequences for the full data set are available on NCBI SRA (BioProject Accession #PRJNA436919). The filtered SNP data set is available from the Dryad Digital Repository https://doi.org/10.5061/dryad.db6177b. R scripts for constrained ordinations and polygenic score calculations are available from the Dryad Digital Repository https://doi.org/10.5061/dryad.b0p66dn.

ORCID

Amanda Xuereb https://orcid.org/0000-0002-3975-2299
Christopher M. Kimber https://orcid.org/0000-0003-4620-0166
Louis Bernatchez https://orcid.org/0000-0002-8085-9709
Marie-Josée Fortin https://orcid.org/0000-0002-9935-1366

REFERENCES

Abraham, J. P., Baringer, M., Bindoff, N. L., Boyer, T., Cheng, L. J., Church, J. A., … Willis, J. K. (2013). A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. *Reviews of Geophysics*, *51*, 450–483.

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Garcia, H. E., … Johnson, D. R. (2010). World Oceans Atlas 2009 Volume 2: Salinity. In S. Levitus (Ed.), *NOAA Atlas NESDIS 69*. Washington, DC, USA: U.S. Government Printing Office.

Babin, C., Gagnaire, P.-A., Pavey, S. A., & Bernatchez, L. (2017). RAD-Seq reveals patterns of additive polygenic variation caused by spatially-varying selection in the Americal eel (*Anguilla rostrata*). *Genome Biology and Evolution*, *9*, 2974–2986.

Bay, R. A., & Palumbi, S. R. (2014). Multilocus adaptation associated with heat resistance in reef-building corals. *Current Biology*, *24*, 2952–2956.

Bay, R. A., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C. K., Lasky, J. R., … Ralph, P. (2017). Predicting responses to contemporary environmental change using evolutionary response architectures. *The American Naturalist*, *189*, 463–473.

Benestan, L., Quinn, B., Laporte, M., Maaroufi, H., Rochette, R., & Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, *25*, 5073–5092.

Benjamini, Y., & Hochberg, Y. (1994). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *1*, 289–300.

Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., & Petrov, D. A. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. *PLoS Genetics*, *10*, e1004775.

Bernatchez, L. (2016). On the maintenance of genetic variation and adaptation to environmental change: Considerations from population genomics in fishes. *Journal of Fish Biology*, *89*, 2519–2556.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., … Withler, R. E. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology and Evolution*, *32*, 665–680.

Bible, J. M., & Sanford, E. (2016). Local adaptation in an estuarine foundation species: Implications for restoration. *Biological Conservation*, *193*, 95–102.

Binyon, J. (1972). *Physiology of echinoderms*. Oxford: Pergamon Press.

Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology*, *89*, 2623–2632.

Bonin, A., Nicole, F., Pompanon, F., Miaud, C., & Taberlet, P. (2007). Population adaptive index: A new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology*, *21*, 697–708.

Bosch, S. (2017). *sdmpredictors*: Species distribution modeling predictor datasets. R Package Version 0.2.6. https://CRAN.R-project.org/package=sdmpredictors.

Bourret, V., Dionne, M., Kent, M. P., Lien, S., & Bernatchez, L. (2013). Landscape genomics in Atlantic Salmon (*Salmo salar*): Searching for gene-environment interactions driving local adaptation. *Evolution*, *67*, 3469–3487.

Bradbury, I. R., Hubert, S., Higgins, B., Borza, T., Bowman, S., Paterson, I. G., … Bentzen, P. (2010). Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proceedings of the Royal Society B*, *277*, 3725–3734.

Breitburg, D. L., Hondorp, D. W., Davias, L. A., & Diaz, R. J. (2009). Hypoxia, nitrogen, and fisheries: Integrating effects across local and global landscapes. *Annual Review of Marine Science*, *1*, 329–349.

Cameron, R. A., Samanta, M., Yuan, A., He, D., & Davidson, E. (2009). SpBase: The sea urchin genome database and web site. *Nucleic Acids Research*, *37*, D750–D754.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetics Research*, *70*, 155–174. https://doi.org/10.1017/S0016672397002954

Conover, D. O. (1998). Local adaptation in marine fishes: Evidence and implications for stock enhancement. *Bulletin of Marine Science*, *62*, 477–493.

De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*, 1383–1399. https://doi.org/10.1111/mec.12182

De Villemereuil, P., Frichot, É., Bazin, É., François, O., & Gaggiotti, O. E. (2014). Genome scan methods against more complex models: When and how much should we trust them? *Molecular Ecology*, *23*, 2006–2019. https://doi.org/10.1111/mec.12705

De Wit, P., & Palumbi, S. R. (2013). Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, *22*, 2884–2897.

DFO. (2014). Interim Advice for the Development of Sea Cucumber (*Parastichopus californicus*) Aquaculture in British Columbia. DFO Canadian Science Advisory Secretariat Science Response 2014/005.

DFO. (2016). Integrated Fisheries Management Plan Summary. Sea Cucumber (*Parastichopus californicus*) By Dive. Pacific Region 2016/2017.

Do Prado, F. D., Vera, M., Hermida, M., Bouza, C., Pardo, B. G., Vilas, R., … Martínez, P. (2018). Parallel evolution and adaptation to environmental factors in a marine flatfish: Implications for fisheries and aquaculture management of the turbot (*Scophthalmus maximus*). *Evolutionary Applications*, *11*, 1322–1341.

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, *196*, 483–493. https://doi.org/10.1016/j.ecolmodel.2006.02.015

Dray, S., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., …Wagner, H. H. (2017). adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.0-9.

Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*, 285–298.

Flanagan, S. P., Forester, B. R., Latch, E. K., Aitken, S. N., & Hoban, S. (2017). Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications*, *11*, 1035–1052.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993.

Foreman, M. G. G., Callendar, W., Masson, D., Morrison, J., & Fine, I. (2014). A model simulation of future oceanic conditions along the British Columbia continental shelf. Part II: Results and analyses. *Atmosphere-ocean*, *52*, 20–38. https://doi.org/10.1080/07055900.2013.873014.

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Molecular Ecology*, *27*, 2215–2233.

Fuentes-Pardo, A. P., & Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations, and practical recommendations. *Molecular Ecology*, *26*, 5369–5406.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, *27*, 489–496.

Gagnaire, P.-A., & Gaggioti, O. E. (2016). Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. *Current Zoology*, *62*, 603–616.

Gagnaire, P.-A., Normandeau, E., Côté, C., Hansen, M. M., & Bernatchez, L. (2012). The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). *Genetics*, *190*, 725–736.

Gobler, C. J., DePasquale, E. L., Griffith, A. W., & Baumann, H. (2014). Hypoxia and acidification have additive and synergistic negative effects on the growth, survival, and metamorphosis of early life stage bivalves. *PLoS ONE*, *9*, e83648.

Golbuu, Y., Gouezo, M., Kurihara, H., Rehm, L., & Wolanski, E. (2016). Long-term isolation and local adaptation in Palau's Nikko Bay help corals thrive in acidic waters. *Coral Reefs*, *35*, 909–918.

Guo, B., DeFaveri, J., Sotelo, G., Nair, A., & Merilä, J. (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biology*, *13*, 19. https://doi.org/10.1186/s12915-015-0130-8

Hamel, J., & Mercier, A. (2008). Population status, fisheries and trade of sea cucumbers in temperate areas of the Northern Hemisphere. In V. Toral-Granda, A. Lovatelli, & M. Vasconcellos (Eds.), *Sea cucumbers. A global review of fisheries and trade. FAO Fisheries and Aquaculture Technical Paper. No. 516* (pp. 257–291). Rome: FAO.

Hecht, B. C., Matala, A. P., Hess, J. E., & Narum, S. R. (2015). Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology*, *24*, 5573–5595.

Hoegh-Guldberg, O., Poloczanska, E. S., Skirving, W., & Dove, S. (2017). Coral reef ecosystems under climate change and ocean acidification. *Frontiers in Marine Science*, *4*, 158. https://doi.org/'0.3389/fmars.2017.00158

Hoegh-Guldberg, O., Cai, R., Poloczanska, E. S., Brewer, P. G., Sundby, S., Hilmi, K., … Jung, S. (2014). The Ocean. In V. R. Barros, C. B. Field, D. J. Dokken, M. D. Mastrandrea, K. J. Mach, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, & L. L. White (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Part B: Regional aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change* (pp. 1655–1731). Cambridge UK and New York NY, USA: Cambridge University Press.

Jeffery, N. W., Stanley, R. R. E., Wringe, B. F., Guijarro-Sabaniel, J., Bourret, V., Bernatchez, L., … Bradbury, I. R. (2017). Range-wide parallel climate-associated genomic clines in Atlantic salmon. *Royal Society Open Science*, *4*, 171394.

Jensen, J. D., Foll, M., & Bernatchez, L. (2016). The past, present and future of genomic scans for selection. *Molecular Ecology*, *25*, 1–4.

Joost, S., Bonin, A., Bruford, M. W., Després, L., Conord, C., Erhardt, G., & Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: Towards a landscape genomics approach to adaptation. *Molecular Ecology*, *16*, 3955–3969.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*, 1225–1241.

Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., … Gattuso, J.-P. (2013). Impacts of ocean acidification on marine organisms: Quantifying sensitivities and interaction with warming. *Global Change Biology*, *19*, 1884–1896.

Kuo, E. S. L., & Sanford, E. (2009). Geographic variation in the upper thermal limits of an intertidal snail: Implications for climate envelope models. *Marine Ecology Progress Series*, *388*, 137–146.

Lambert, P. (1997). *Sea cucumbers of British Columbia, Southeast Alaska and Puget Sound*. Vancouver, BC: UBC Press.

Laporte, M., Pavey, S. A., Rougeux, C., Pierron, F., Lauzent, M., Budzinski, H., … Bernatchez, L. (2016). RAD-sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*, *25*, 219–237.

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*, 271–280.

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology*, *17*, 142–152.

Lundgren, P., Vera, J. C., Peplow, L., Manel, S., & van Oppen, M. J. H. (2013). Genotype-environment correlations in corals from the Great Barrier Reef. *BMC Genetics*, *14*, 9. https://doi.org/10.1186/1471-2156-14-9

Manel, S., Perrier, C., Pratlong, M., Abi-Rached, L., Paganini, J., Pontarotti, P., & Aurelle, D. (2016). Genomic resources and their influence on the detection of the signal of positive selection in genome scans. *Molecular Ecology*, *25*, 170–184.

Matear, R. J., Hirst, A. C., & McNeil, B. I. (2000G). Changes in dissolved oxygen in the Southern Ocean with climate change. *Geochemistry, Geophysics, Geosystems*, *1*, 2000GC000086.

Morrison, J., Callendar, W., Foreman, M. G. G., Masson, D., & Fine, I. (2014). A model simulation of future oceanic conditions along the British Columbia continental shelf. Part I: Forcing fields and initial conditions. *Atmosphere - Ocean*, *52*, 1–19.

Munday, P. L., Donelson, J. R. M., & Domingos, J. A. (2017). Potential for adaptation to climate change in a coral reef fish. *Global Change Biology*, *23*, 307–317.

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*, 184–194.

Nielsen, E. S., Beger, M., Henriques, R., Selkoe, K. A., & von der Heyden, S. (2017). Multispecies genetic objectives in spatial conservation planning. *Conservation Biology*, *31*, 872–882.

O'Connor, M. I., Bruno, J. F., Gaines, S. D., Halpern, B. S., Lester, S. E., Kinlan, B. P., & Weiss, J. M. (2007). Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences*, *104*, 1266–1271.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., ...Wagner, H. H. (2017). vegan: Community Ecology Package. R package version 2.4-5. https://CRAN.R-project.org/package=vegan.

Osovitz, C. J., & Hoffman, G. E. (2005). Thermal history-dependent expression of the *hsp70* gene in purple sea urchins: Biogeographic patterns and the effect of temperature acclimation. *Journal of Experimental Biology and Ecology*, *327*, 134–143.

Palumbi, S. R., Barshis, D. J., Traylor-Knowles, N., & Bay, R. A. (2014). Mechanisms of reef coral resistance to future climate change. *Science*, *344*, 895–898.

Pardo, L. M., & Johnson, L. E. (2005). Explaining variation in life-history traits: Growth rate, size, and fecundity in a marine snail across an environmental gradient lacking predators. *Marine Ecology Progress Series*, *296*, 229–239.

Pespeni, M. H., & Palumbi, S. R. (2013). Signals of selection in outlier loci in a widely dispersing species across an environmental mosaic. *Molecular Ecology*, *22*, 3580–3597.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, *7*, e32253.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*, 4348–4370.

Rocha-Olivares, A., & Vetter, R. D. (1999). Effects of oceanographic circulation on the gene flow, genetic structure, and phylogeography of the rosethorn rockfish (*Sebastes helvomaculatus*). *Canadian Journal of Fisheries and Aquatic Sciences*, *56*, 803–813.

Russell, M. (2013). Echinoderm responses to variation in salinity. *Advances in Marine Biology*, *66*, 171–212.

Sanford, E., & Kelly, M. (2011). Local adaptation in marine invertebrates. *Annual Review of Marine Science*, *3*, 509–535.

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *11*, 807–820.

Sbrocco, E. J., & Barber, P. H. (2013). MARSPEC: Ocean climate layers for marine spatial ecology. *Ecology*, *94*, 979. https://doi.org/10.1890/12-1358.1

Schoville, S. D., Bonin, A., François, O., Lobreaux, S., Melodelima, C., & Manel, S. (2012). Adaptive genetic variation on the landscape: Methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, *43*, 23–43.

Sea Urchin Genome Sequencing Consortium. (2006). The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science*, *314*, 941–952. https://doi.org/10.1126/science.1133609

Sgrò, C. M., Lowe, A. J., & Hoffmann, A. A. (2011). Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, *4*, 326–337.

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., & Neale, D. B. (2013). Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, *9*, 901–911.

Stanley, R. E., DiBacco, C., Lowen, B., Beiko, R. G., Jeffery, N. W., van Wyngaarden, M., ... Bradbury, I. R. (2018). A climate-associated multispecies cryptic cline in the northwest Atlantic. Science. *Advances*, *4*, eaaq0929.

Stickle, W. B., & Diehl, W. J. (1987). Effects of salinity on echinoderms. *Echinoderm Studies*, *2*, 235–285.

Sunday, J. M., Popovic, I., Palen, W. J., Foreman, M. G. G., & Hart, M. W. (2014). Ocean circulation model predicts high genetic structure observed in a long-lived pelagic developer. *Molecular Ecology*, *23*, 5036–5047.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, *21*, 272–281.

van Wyngaarden, M., Snelgrove, P. V. R., DiBaccio, C., Hamilton, L. C., Rodrguez-Ezpeleta, N., Jeffrey, N. W., ... Bradbury, I. R. (2017). Identifying patterns of dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*, using RADseq derived SNPs. *Evolutionary Applications*, *10*, 102–117.

von der Heyden, S. (2017). Making evolutionary history count: Biodiversity planning for coral reef fishes and the conservation of evolutionary processes. *Coral Reefs*, *36*, 183–194.

Xuereb, A., Benestan, L., Normandeau, E., Curtis, J. M. R., Bernatchez, L., & Fortin, M.-J. (2018). Asymmetric oceanographic processes mediate connectivity and population genetic structure as revealed by RADseq in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Molecular Ecology*, *27*, 2347–2364.

Yanick, J. F., Heath, J. W., & Heath, D. D. (2003). Survival and growth of local and transplanted blue mussels (*Mytilus trossulus*, Lamark). *Aquaculture Research*, *34*, 869–875.

Yeaman, S., & Otto, S. P. (2011). Establishment and maintenance of adaptive genetic divergence under mutation, selection, and drift. *Evolution*, *65*, 2123–2129.

Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., ... Aitken, S. N. (2016). Convergent local adaptation to climate in distantly related conifers. *Science*, *353*, 1431–1433.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Xuereb A, Kimber CM, Curtis JMR, Bernatchez L, Fortin M-J. Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data. *Mol Ecol*. 2018;27:5035–5048. https://doi.org/10.1111/mec.14942